

DOCUMENT RESUME

ED 355 380

CE 063 214

AUTHOR Hater, John J.
TITLE Work Keys: Integrating Work Keys in the Business Environment.
PUB DATE 14 Aug 92
NOTE 14p.; Paper presented at the Annual Convention of the American Psychological Association (100th, Washington, DC, August 14, 1992).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Descriptive (141)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Adult Education; Basic Skills; Construct Validity; *Content Validity; Job Analysis; *Job Skills; Labor Force Development; Listening Skills; Occupational Tests; Reading Skills; Service Occupations; *Service Workers; Task Analysis; *Test Content; Testing; Test Items; *Test Use; Vocational Education; Writing Tests
IDENTIFIERS Federal Express; *Work Keys (ACT)

ABSTRACT

Work Keys (occupational tests developed by American College Testing) could support an employer's human resource function in a number of ways: (1) communicating to educators the skill requirements for an employer's particular jobs on a national basis; (2) providing students with a realistic preview of skills needed for jobs and an assessment of their standing on those skills; (3) helping employers selection decisions; and (4) helping employers in need of upgrading their current work force. For most employers to use Work Keys assessments for selection decisions, the question is whether a content validation strategy will be adequate. A content strategy was used to validate the Reading for Information and Listening and Writing tests for the customer service agent job at Federal Express. Information from observations of and interviews with job incumbents and managers was used to develop task statements, ability statements, and task clusters. Fifty-eight task statements were put into a questionnaire that asked for two ratings for each task: how often it is performed and its importance. The ability questionnaire with 26 ability statements asked for 2 ratings from instructors: importance and when the ability was acquired. Linkages that were made between tasks and abilities were inferences concerning whether an ability was a prerequisite for performing the task. A Content Validity Panel of eight subject matter experts (SMEs) took the Work Keys tests. Data on response fidelity were gathered. All SMEs thought the tests were fair and valid measures of the skills used in the job. (YLB)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED355380

WORK KEYS:
**INTEGRATING WORK KEYS IN THE BUSINESS
ENVIRONMENT**

JOHN J. HATER
FEDERAL EXPRESS CORPORATION

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☒ This document has been reproduced as
received from the person or organization
originating it.
☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official
OERI position or policy.

PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

J. J. Hater

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

**IN J. D. WEST (CHAIR), WORK KEYS--SUPPORTING THE
TRANSITION FROM SCHOOL TO WORK. SYMPOSIUM AT THE ONE
HUNDREDTH ANNUAL CONVENTION OF THE AMERICAN PSYCHOLOGICAL
ASSOCIATION, AUGUST 14, 1992, WASHINGTON, D. C.**

CF 063214

Work Keys:
Integrating Work Keys in the Business Environment

How Work Keys Could Support An Employer's Human Resource Function

Work Keys could support an employer's human resource function in a number of ways, unlike any other assessment effort. First, it could communicate to educators the skill requirements for an employer's particular jobs on a national basis. This is particularly advantageous for national employers who find it impossible to develop close associations with school entities across the nation. With Work Keys, the ties between businesses and schools could move beyond well meaning but limited adopt a school efforts. As the required skills are communicated to educators via job profiles, we could see changes to curriculum and even how subjects are taught.

For example, the Work Keys Listening and Writing test represents one skill area that Federal Express needs for its Customer Service Agents. However, we know of no high school or vocational school which teaches combined listening and writing skills. Educators who had students wishing to work for Federal Express as a Customer Service Agent (an entry level position), would be placing their students in a favorable position for securing a job with Federal Express if they had a course in listening and writing.

As Work Keys communicates to educators what are the job specific required skills for different employers, it would also be providing students with a realistic preview of the skills needed for jobs and an assessment of their standing on these skills. This is the second function of Work Keys, one which would establish for students a direct link between what they learn in school and what is required to get a job. If Work Keys assessments are given well before high school graduation, then it would allow students time to acquire the needed skills.

A third and the most obvious function of Work Keys is that it could be used for selection decisions. How it is used for selection decisions could run along a continuum ranging from a stated preference for the skill by an employer to using it as a gate or a minimum requirement prior to consideration in other parts of the selection process.

And as part of the process of making selections, Work Keys could provide feedback to applicants concerning the reasons for not being hired by a particular employer. Again, if secondary or vocational schools have curriculum addressing the skills, it would give applicants an avenue to improve the deficient skills.

This function of Work Keys could do a lot to improve the reputation of pre-employment testing through examinee empowerment. That is, what options do applicants have when they are turned down for a job because they did not score above the minimum on aptitude tests? One option that they can use and that can be very costly to employers is filing a lawsuit to challenge the validity or business necessity of testing practices. It would seem that an applicant would be less likely to file a lawsuit if there had been years of advance notice (during school) that a skill would be required or if there was a clear avenue for obtaining

the prerequisite skill in the future. It would also seem that a court or a jury would be less likely to find for a plaintiff when a complaint concerned a criterion-referenced test and when there was remedial training available in the community to improve the required skill.

A fourth function of Work Keys is that it could be used by employers who are in need of upgrading their current work force to accommodate new technology, more sophisticated manufacturing processes, etc. For example, many employers must assess training needs and provide remedial training or education before the introduction of new technology to the current work force. A related function would be using Work Keys assessments to diagnose performance problems found in a particular employee. If schools have curriculum addressing the deficient skill for the employee with the performance problem, a termination action could be avoided.

Just as students taking the ACT assessment for college admissions can give ACT permission to release their scores to colleges, students taking Work Keys assessments will also be able to give permission to ACT to release their scores to employers. Employers could then go to ACT and request a recruitment pool for particular jobs by skills and locations. This would be especially advantageous if an employer was expanding and trying to find a location with a large available pool of qualified applicants.

How Can Employers Validate Work Keys Assessments for Selection

Of course, many of these functions and benefits of Work Keys will go unrealized if employers cannot validate the assessments for selection decisions. Selection decisions represent the bottom line and educators and students will not take Work Keys skills very seriously if employers only say that they have jobs that use the skills.

For most employers to use Work Keys assessments for selection decisions, the question will become whether or not a content validation strategy will be adequate for Work Keys assessments because it generally will not be possible to conduct criterion-related or construct validation studies due to costs or a lack of adequate samples. In the more typical application of the content validation strategy, the goal is to simulate tasks in tests for a specific job. With Work Keys, content validity would be used to support one test across many jobs that may not share identical tasks. This is the same problem for any off the shelf test that was content validated and used for selection decisions across jobs. We are hoping the problem will not be insurmountable for Work Keys tests because they are multi-level and criterion-referenced and are being developed with close consultation with employers.

The rest of this paper will describe our efforts to use a content strategy to validate the Reading for Information and Listening and Writing tests for the Customer Service Agent job at Federal Express. We are very fortunate to have Dr. Goldstein as our discussant because he is a leading expert on the content validity process. We will be using terms and ideas extensively from his chapter on content validity in the next Frontier Series.

The Customer Service Agent Job

There are over 2,000 Federal Express Customer Service Agents working at 15 call centers located across the United States. About 70% are full-time and 30% are part-time. It is an entry level position and all new hires receive five weeks of paid full-time training. The average age is 34 years and the average tenure with the company is 3.5 years.

This customer contact job is conducted almost entirely on the telephone in front of a computer terminal in which much of the required information is found and various transactions are made. Agents also consult various written hard copy resources for information. The major purposes of the position are to provide information about services, create and update customer files, determine the status or location of packages, initiate a request for package pickup, and document and solve special requests, problems, and complaints. To minimize having to transfer a customer to a more specialized department, agents are repeatedly trained and equipped to deal with a wide variety of requests and problems.

Job Analysis

Job analysis of the Customer Service Agent job was being conducted by Dr. David Brooks and myself as part of another validation study with four other jobs when we were introduced to Work Keys. This job analysis information was used as part of the Work Keys validation. Over 35 job incumbents and 20 first line managers were observed and/or interviewed (individually or in groups) at seven of the call centers. Job incumbents were used primarily for the development of task statements and managers were used primarily for development of ability statements. Task clusters were also developed during these interviews.

Task questionnaires. The final list of 58 task statements were put into a questionnaire that asked for two ratings for each task: (a) how often performed? (response options were 0=never, 1=infrequently, 2=less frequently, 3=moderately frequent, 4=more frequent, and 5=most frequent), and (b) importance (response options were 1=of no importance, 2=of minor importance, 3=of importance, 4=of high importance, and 5=of critical importance). Managers administered the questionnaires to 200 job incumbents at the 15 centers. A cover memo explained that the questionnaire information would be used for the purpose of developing pre-employment tests. A total of 197 usable questionnaires were returned to us. Comparisons of the questionnaire sample and the total agent population showed the sample was representative of the population. For example, 45% of the sample was minority and 40% of the population is minority, and the average overall performance rating of the sample was 6.3 and the average overall performance rating of the population was 6.3 (on a scale of 1.0 to 7.0).

Tasks were eliminated if they were performed by less than 70% of the total sample (i.e., more than 30% chose "never" on the frequency scale). Tasks were also eliminated if the average importance was less than 3.0 among the total sample. No tasks were eliminated based on the frequency criterion, and two tasks were eliminated due to the importance criterion.

Ability Questionnaire. The ability questionnaire consisted of 26 ability statements written for the job analysis of the agent job and the other jobs in the larger validation study. Two ratings were included in this questionnaire: (a) importance (response options were 1=of no importance, 2=of minor importance, 3=of importance, 4=of high importance, and 5=of critical importance), and (b) when acquired? (response options were B=before and A=after being hired/promoted into this job). This questionnaire was administered in a face to face meeting with 15 instructors from the various centers. These instructors are responsible for delivering the new hire five week training and they seemed to be in the best position to judge when an ability was normally acquired. Although the questionnaires were administered face to face, responses were made individually after we read each ability statement and then elaborated on the ability through examples.

An ability was retained in the job analysis if 13 of the 15 raters agreed that the ability was acquired before hiring/promotion. The criterion of 13 was determined by using Lawshe's (1975) content validity ratio. There were eight of the 26 abilities that were eliminated due to this criterion. All of the surviving abilities had average importance ratings above 3.0.

Linkages. Linkages between tasks and abilities were initially made by the psychologists working on the project. These linkages were then confirmed or deleted by a panel of four first line managers and two instructors moderated by a psychologist. The linkages were based on a consensus of the panel who also added linkages to those initially made. The linkages were inferences concerning whether an ability is a prerequisite for performing a particular task. The linkages were made not only for individual task statements but also for task clusters.

The following lists the abilities that are related to the Work Keys assessments of Reading for Information and Listening and Writing, the number of tasks linked to the ability, and some exemplar linkages.

1. Reading for Information. Ability to identify and understand information presented in materials such as directions, instructions, memos, newsletters, and guides.

This ability was linked to 31 of the 56 task statements and to all of the five task clusters. Exemplar linkages included the following tasks:

1. Read for understanding FedEx daily information on PF4/DRS Data Reference System and information in SRG's, Service Guide, implementation releases, job aids, memos, and E-Mail.
4. Find information in print resources (SRG, Service Guide, agent resource manual, Zip code directory, implementation releases, job aids) to inform customers about products, services, restrictions, procedures, etc.
11. Give directions to FedEx facility, hours of operation, or services available to a customer by reading information for a particular Zip code or four letter ID on the PF6 FedEx Location Information screen.
27. Determine whether or not a customer should have an account number by searching for exception or special instructions in the PF4 Data Reference System (Category 12).

2. Listening. Ability to receive information by hearing alone (audio tape or telephone presentation). It includes hearing accurately and confirming understanding.

This ability was linked to 42 of the 56 task statements and to all of the five task clusters. Exemplar linkages included the following tasks:

5. Assist a customer in choosing FedEx services by asking fact-finding questions and matching customer's needs with the appropriate services.
30. Verify the spelling of names or streets to ensure accuracy of records.
45. Update a Customer Exception Request (CER) for a customer with a typed-in summary of the facts communicated by a customer.
53. Identify and qualify sales leads and send messages to account executives using the Customer Exception Request (CER) screen.

3. Taking Down Orally Presented Information. Ability to write accurately numbers, letters, names, or addresses presented orally.

This ability was linked to 39 tasks and to all of the five task clusters.

Exemplar linkages included the following:

41. Initiate a Customer Exception Request (CER) by entering shipper, recipient, package, and caller information, service type, priority situation, appropriate queue type, and comments.
45. Update a Customer Exception Request (CER) for a customer with a typed-in summary of the facts communicated by a customer.
50. Update customer complaints and compliments (COV) with typed-in comments.
51. Compose and type Y messages to other locations (messages to station and dispatch employees, etc.).

It should also be noted that the ability, Composing Written Communications, survived the ability analyses but it was not formally linked to tasks by the panel because it was thought that developing a Composing Written Communications assessment was beyond the resources of the larger validation project at Federal Express. Composing Written Communications was defined as the ability to compose clear and understandable communications in writing such as notes and messages. The sample of tasks presented above that are linked to Taking Down Orally Presented Information as well as other tasks could have been linked to this ability.

In-house Test Vs. Work Keys Test

With the abilities linked to the job tasks, we were in a position to develop tests simulating the tasks that would demonstrate the abilities. We have begun our own test development, but at the same time we were contacted by ACT about the Work Keys assessments. However, by using the Work Keys tests we knew that these tests would not approximate the job tasks as much as tests developed in house (i.e., lower physical fidelity). But as Dr. Goldstein stated in his chapter, the psychological fidelity of a test rather than its physical fidelity is of greater importance from the standpoint of content validity. That is, do the Work Keys

tests simulate the physical components of the job such that the abilities required to perform the tasks are called forward or displayed?

Dr. Goldstein stated that there are three important issues that contribute to the psychological fidelity and thereby, the content validity of a test, and we will evaluate the Work Keys tests for the Customer Service Agent job in light of them. They are:

1. The degree to which the ability required by a job is tapped in a test, regardless of the testing mode.
2. Whether the testing mode is representative of the way tasks are accomplished on the job.
3. The degree to which the test is scored for knowledge, skill, or ability not even required by the job.

A related issue with the Work Keys tests concerns what level of a test has psychological fidelity with the job. With Work Keys tests having multiple levels, we had to collect data to advise us on what is the valid level of a test for our job.

Content Validity Panel

The content validity panel consisted of eight subject matter experts (SME's): two first line managers and six Sr. Customer Service Agents. The panel was conducted by Dr. Joy McLarty, Dr. Tim Vansickle, and myself. We began the session with an overview of Work Keys and a summary of our validation efforts to date. The SME's were provided copies of the task statements and the linkages between the tasks and the abilities of interest for Work Keys: Reading for Information, Listening, and Taking Down Orally Presented Information. We then conducted a discussion of the how these abilities are used on the job using the linkage documentation.

Reading for Information Test. After the discussion, the SME's took abbreviated versions of the Reading for Information test. The correct answers to the test were reviewed when they finished. To gather data on the stimulus fidelity of the test, SME's were next given 15 reading passages that included three passages for each of the five test levels ordered randomly. After reading each passage, a SME individually judged whether the passage was easier, about the same difficulty, or more difficult than those typically read on the job.

To gather data on the response fidelity of the test, each SME individually judged the difficulty of questions for the reading passages. Each SME was given 10 reading passages and its questions for one of the five levels. We gave pairs of SME's level four, five, six, or seven. (Our original intent was to give a SME the level that he/she said was about the same difficulty for the job, but each SME did not consistently judge one particular level as about the same difficulty.) For each set of questions under a passage, the SME judged whether the questions were easier, about the same, or more difficult than those typically answered from reading material on the job. Then using the Reading for Information level descriptions and example passages, each SME made an overall judgment for the level of reading that is typically required by the job.

To help answer whether the test tapped the Reading for Information ability, each SME answered the following question:

"Reading for Information was an ability identified as essential for the Customer Service Agent job by subject matter experts. Consider to what extent the reading passages and questions like those you said were required for this job provide an opportunity for a person to demonstrate this ability. Rate whether the ability can be demonstrated to a great extent, to moderate degree, to slight degree, or not at all, using this test."

Lastly, each SME answered yes or no to the question of whether a test composed of passages and questions like those he/she said were required for the job would be a fair and valid measure of the reading skills used in the job.

Results of Reading for Information Judgments. Presented below are the judgments for the difficulty of the reading passages. These judgments show that level 3 and 4 were easier or about the same level as the job requirements and 79% of the judgments about level 5 passages were easier or same as the job requirements.

Difficulty of Reading Passages: Judgments of Eight SME's for Three Passages from Each Level

Level	Easier	Same	More Difficult
3	21	3	0
4	14	8	1
5	11	8	5
6	6	10	8
7	4	12	8

The 80 judgments on the difficulty of the reading questions at level 4 through 7 were all easier or the same as the job requirements except for three more difficult judgments. The overall judgments using the level descriptions were one for level 4, three for level 5, and four for level 6. All of the eight judges indicated that the ability could be demonstrated to a moderate degree using this test and all judges indicated that the test would be a fair and valid measure of the reading skills used in the job.

Listening and Writing Test. Prior to making any judgments about the Listening and Writing test, the SME's took an abbreviated version of the test. To judge test stimulus fidelity, SME's listened to 12 messages that included three messages from each of the four levels randomly ordered, and then each SME judged whether the message was easier, about the same difficulty, or more difficult than those typically listened to and written up on the job.

To gauge the response fidelity of the listening score on the test, each SME considered the listening scoring criteria (levels 0 through 5) and judged the level of listening typically required by the job, regardless of the length and complexity of the message. Each SME also judged the extent to which listening could be demonstrated by the test and answered yes or no to whether the test would be a fair and valid measure of the listening skills used in the job.

To gauge the response fidelity of the writing score on the test, each SME also considered the writing scoring criteria (levels 0 through 5) and judged the level of writing typically required, regardless of the length and complexity of the

message. As with the other tests, each SME judged the extent to which writing ability could be demonstrated by the test and answered yes or no to the question about fairness and validity.

Results of Listening and Writing Judgments. Presented below are the judgments for the difficulty of the taped messages. All of the messages at levels 1 and 2 were judged easier or the same. Only two judgments were more difficult for level 3 messages and level 4 messages, and level 3 received twice as many easier judgments than level 4.

Difficulty of Messages: Judgments of Eight SME's for Three Messages from Each Level

Level	Easier	Same	More Difficult
1	17	7	0
2	11	13	0
3	6	16	2
4	3	19	2

Considering the Listening scoring criteria, there was one judgment for level 4 and seven judgments for level 5 as the level of listening typically required by the job. Most of the judges (six) indicated that listening ability could be demonstrated to a great extent using the test and two of the judges said to a moderate degree. All of the SME's said that the test would be a fair and valid measure of the listening skills used on the job.

For the Writing scoring criteria, there were two judgments for level 3, three judgments for level 4, and three judgments for level 5. On the question of could writing ability be demonstrated by the tests, four of the judges indicated to a great extent and four of the judges indicated to a moderate degree. Again, all of the SME's indicated that the test would be a fair and valid measure of the writing skills used on the job.

Discussion

To even the casual observer of the Customer Service Agent job, reading, listening, and writing are major parts of the job. But do the Work Keys tests have enough content validity to be used for selection decisions? To answer this question, we will discuss the data we have collected and our observations in light of the issues identified by Dr. Goldstein.

Reading for Information Test. To answer the question of whether the reading for information ability was tapped by the test, we found that all of the eight SME's said the ability could be demonstrated to a moderate degree using the test.

Concerning the question of whether the test mode or procedure was representative of the job, time did not permit asking the SME's whether each of the 31 tasks linked to reading was represented in the test. However, even without their corroborating judgments, it is obvious that agents read prose documents and answer questions for customers. Customers ask questions orally and agents answer the questions orally with the aid of written materials, but

simulating such an oral exchange would be unfeasible and would not seem to contribute to the measurement of reading ability, but rather oral communication.

Since representativeness also concerns the difficulty levels of the reading passages and questions, the judgments about the difficulty of the sample passages and the overall judgment using the level descriptions were helpful. I would set the minimum requirement at level 5 based on these judgments. A readability analysis of the passages would also be helpful for determining their match with the job because we could compare it to a reading analysis made already of the materials used in the job. Readability scales might have problems (Burrill, 1987), but we would not use it as the only criterion to judge passage difficulty.

Looking at the Reading for Information level 5 description concerning the questions, it states that the questions typically call for applying information given in a passage to a situation that is not specifically described in the passage. This is representative of the agent job because they often have to read and apply a written policy, procedure, etc. to a customer's specific situation. For example, a task linked to this ability was: "Provide delivery commitment information to a customer based on type of service, package dimensions, origin/destination Zip code, and contents (if dangerous goods)." Moreover, their job is more than answering questions using answers paraphrased from written materials (level 4), but it is less than answering questions using information that is not explicitly stated in writing (level 6), at least at the entry level to the job.

The last issue to consider for the Reading for Information test is whether the test is scored for knowledge, skill, or ability not required by the job. What comes to mind here is that prior knowledge about a particular subject or cultural familiarity might give some examinees an advantage. I have not examined all of the passages in enough detail to have an opinion about such a potential threat to validity.

Listening and Writing Test. I will first discuss the messages in this test because both the listening and writing scoring criteria use the same messages. Concerning representativeness, only two of the 24 judgments about level 3 passages and only two judgments about level 4 passages were that they were more difficult than the job. Unknown to the judges, level 3 messages contain 14 pieces of information and level 4 messages contain 17 pieces of information. It is not unusual for an agent to handle more than 17 pieces of information from a customer to complete all of the fields on a computer screen, and if an agent has to document in writing a customer request, complaint, or compliment, there is no limit on the amount of information that must be included. It appears then that the level 4 messages are representative for the job.

Listening. Concerning the issue of whether listening ability was tapped by the test, six judges said the ability could be demonstrated to a great extent and two said said to a moderate degree using the test. Regarding representativeness, it is obvious from the job analysis that agents listen to a variety of customer requests, situations, and predicaments for which they must either solve or summarize in writing using notes if they wish. The unrepresentativeness would be that the listening on the job can be supplemented by asking the customer questions, although asking a customer to repeat information already given is considered poor customer service and delaying the call unnecessarily. Since the

message is said twice during the test, this aspect of unrepresentativeness is probably of no consequence. In addition, the test writing task is simply summarization and not problem solving for which the opportunity for asking a customer questions would be appropriate.

The judgments for the listening scoring were seven judgments for level 5 and one for level 4. The definition of level 5 says: "All information is correct. All critical information is present and correct. The message conveys insight into the situation through tone and/or subtle details." This stands in contrast to level 4's definition which states: "... may be missing subtle details or tone ..." I was not surprised that this level was chosen by seven of the judges given the Federal Express culture of achieving 100% service and that leaving out even one subtle detail about a customer's request or problem could contribute to a service failure or customer dissatisfaction. I would conclude therefore that the test is not scored for knowledge, skill, or ability not required by the job unless there is a bias for knowledge of particular subjects or cultural familiarity with content in the messages.

Writing. On the issue of whether writing ability was tapped by the test, four judges said the ability could be demonstrated to a great extent and four judges said the ability could be demonstrated to a moderate degree. On the issue of representativeness, we have examples of messages that agents wrote or rather typed on the job that describe various situations. A related comment made by the SME's after taking this test was that it was very tiresome to write the messages because they type all of their messages into the computer. However, most job applicants would probably be more comfortable with writing their test answers rather than typing them even though typing is more similar to the job.

The judgments for the writing scoring criteria were spread out: two for level 3, three for level 4, and three for level 5. Comparing the definition of level 3 with exemplar messages from agents and the performance review form for this job, I would recommend level 3 as the appropriate level. Level 3 states: "Message conveyed clearly; acceptable for business setting; some mechanical errors; most sentences are complete." Level 4 states: "Message conveyed clearly; appropriate for business setting; may have a few minor mechanical errors; all sentences are complete; may have choppy style." Looking at the exemplar messages from the job it is apparent that not all sentences are complete. In addition, the relevant performance review verbiage only says "providing accurate and complete information ..." and there is no mention of the mechanics of writing or sentence structure. Using level 3, I would conclude therefore that the test is not scored for knowledge, skill, or ability not required by the job unless again there is a bias for knowledge of particular subjects or cultural familiarity with the content in the messages.

Conclusions

The last question asked about each test was whether the SME thought it was a fair and valid measure of the skills used in the job. It was encouraging that all of the SME's thought that all of these tests were fair and valid for the job with the availability of jury trials for unfair treatment complaints!

But what are some of the shortcomings of this analysis that threaten the content validity evidence. First, one content validity panel is not adequate since we are dealing with qualitative judgments and replication would make up for the small number of SME's involved in these judgments. Second, we would have to actually obtain the linkages between Composing Written Communications and the tasks in addition to the linkages of Taking Down Orally Presented Information. Third, according to Dr. Goldstein's chapter we should have obtained the linkages between tasks and abilities through individual independent responses instead of using group consensus data. Fourth, additional judgments by SME's concerning whether particular tasks are represented in the test would be advantageous to corroborate my observations and other data about pieces of information handled on the job, etc.

At the beginning of this paper I said that if Work Keys tests were used for selection decisions, the uses could run along a continuum of an employer's stated preference for the skill to using the tests as gates or minimum criteria. Given the data we have collected at this point, I think we could begin to use these tests at the lowest level of this continuum. That is, I believe there is enough validity evidence to say that Federal Express has a preference for Customer Service Agent applicants with Work Keys Skills of Reading for Information at level 5, Listening at level 5, and Writing at level 3. However, I do not think we have enough validity evidence to say that these are absolute qualifiers or that all applicants must have these skill levels to be considered. Such an application would require fixing the content validity shortcomings and possibly collecting data that showed applicants without these skill levels failed the initial training program or failed on the job.

References

Burrill, L. E. (1987). How well should a high school graduate read? NASSP Bulletin, 71, 61-72.

Goldstein, I. L., Zedeck, S., & Schneider, B. (in press). An exploration of the job analysis-content validity process. In I. L. Goldstein (Ed.), Frontiers Series. Society for Industrial and Organizational Psychology.

Lawshe, C. H. (1975). A quantitative approach to content validity. Personnel Psychology, 28, 563-575.